ARMY RESEARCH LABORATORY

Use of a Correlation Coefficient for Conditional Averaging

by Richard B. Loucks

ARL-TN-91

April 1997

Army Research Laboratory
Adelphi, MD 20783-1197

ARL-TN-91 **April 1997**

Use of a Correlation Coefficient for Conditional Averaging

Richard B. Loucks
Information Sciences and Technology Directorate

Abstract

A method of collecting ensembles for conditional averaging is presented that uses data collected from a plane mixing layer. The correlation coefficient of a sine function and a reference signal were used to determine the adequacy of blocks of the reference signal and the phase alignment for the data. Selection of the sine function period and a correlation coefficient threshold are discussed. Also examined are the effects of the period and threshold level on the number of ensembles captured for inclusion for conditional averaging. Both the selection of threshold correlation coefficient and the choice of data set length determine bandpass filtering; this study examined the effect of these choices on the ability to discriminate reference velocity ensembles for conditional averaging.

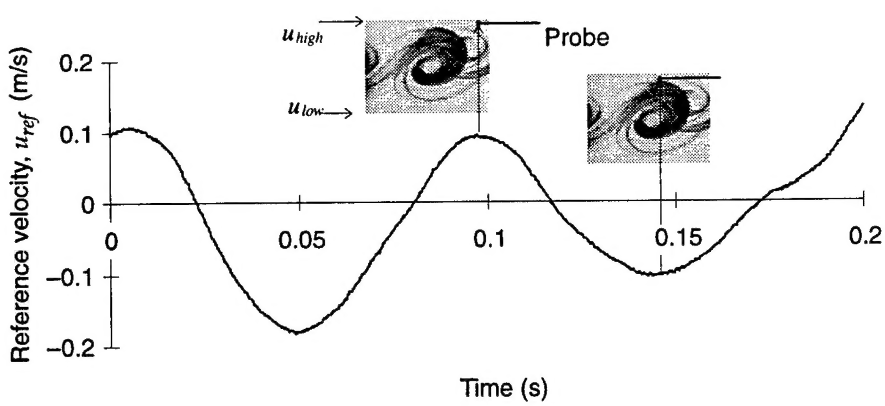
Contents

Figures

1. Introduction

One powerful tool for investigating coherent structures in a turbulent fluid flow is a conditional average resulting from the collection of several conditional samples that are subjected to some averaging technique. The purpose is to determine, given enough events, what properties of turbulent fluid flow are consistent, and to possibly identify coherent structures in an otherwise chaotic flow. To perform a conditional average, one must identify an event in the flow, and some feature about the event with which to align the data to be collected. An example of an event is the passage of a shedding vortex off a bluff body. In this example, the data surrounding the event would be aligned with the point in time at which the velocity is at a maximum when a vortex passes. Once an event is identified and a means of alignment provided, several separate events can be measured and pooled for conditional averaging and evaluation. Subramanian et al[1] describe several techniques for detecting and averaging data based on discriminable events. Davies[2] used peak velocity as a means to detect the passage of shedding vortices in a Von Kármán street, which have a strongly periodic flow. In the case of the turbulent mixing layer, another strongly periodic flow, it is obvious from inspection that the passage of a roller is an excellent candidate for an event. In fact, Hussain and Zaman[3] showed that a single sensor probe, placed several momentum thicknesses from the mixing layer centerline, resulted in a streamwise velocity time series that exhibited clear, periodic events. As in the mixing layer depicted in figure 1, the single hot-wire probe that is located at the high-speed edge of the roller senses a decrease in local velocity induced by the retarding motion of the passing roller. When the roller passes, the local streamwise velocity increases above that of the free stream. The velocity oscillates, and appears nearly sinusoidal with the passage of time.

Figure 1. Passage of roller with respect to reference probe.



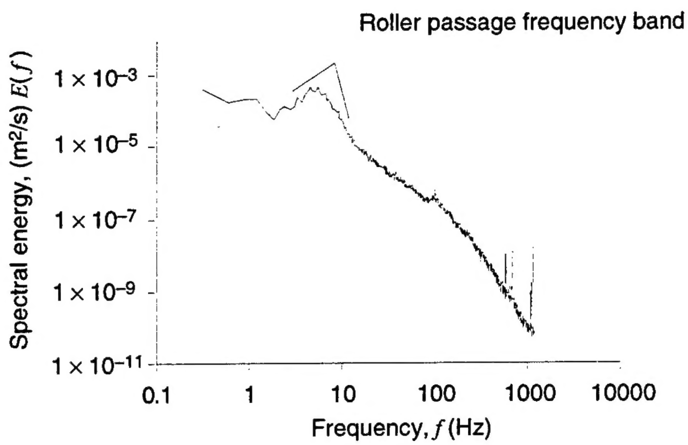
[1]C. S. Subramanian, S. Rajagopalan, R. A. Antonia, and A. J. Chambers, "Comparison of conditional sampling and averaging techniques in a turbulent boundary layer," J. Fluid Mech. **123** (1982), 335–362.

[2]M. E. Davies, "A comparison of the wake structure of a stationary and oscillating bluff body, using a conditional averaging technique," J. Fluid Mech. **75** (1976), 209–231.

[3]A.K.F.M. Hussain and K.B.M.Q. Zaman, "An experimental study of organized motion in the turbulent plane mixing layer," J. Fluid Mech. **159** (1985), 85–104.

We can take advantage of this sinusoid-like signal to detect the passage of the roller, and collect the data of interest surrounding the event. We would align the data merely by identifying some points in the signal, such as the maxima or minima. Unfortunately, the velocity data are not exactly a sine function. In fact, as seen in the fast Fourier transform (FFT) of figure 2, there is a range of frequencies associated with roller passage. An alternative method for filtering the reference velocity and aligning the data is needed.

Figure 2. Fast Fourier transform (FFT) of reference velocity power spectrum, experiment 1.



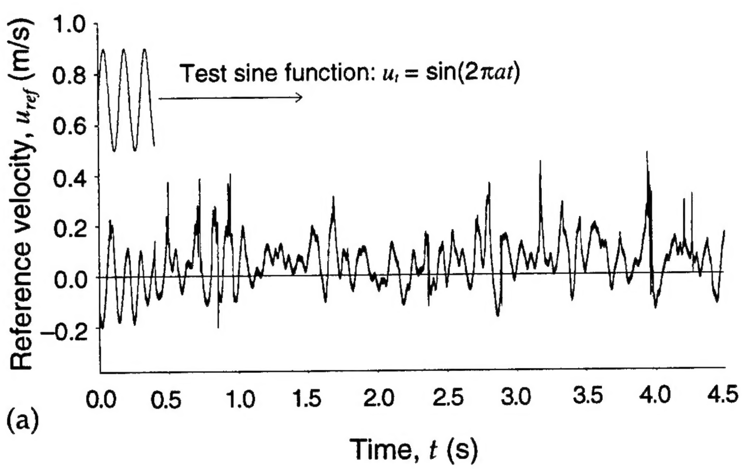
2. Event Detection

One method to detect the passage of a particular type of roller is to use a correlation of the reference velocity with a sine function, as depicted in figure 3a. The correlation of the data with a test sine function results in a correlation coefficient:

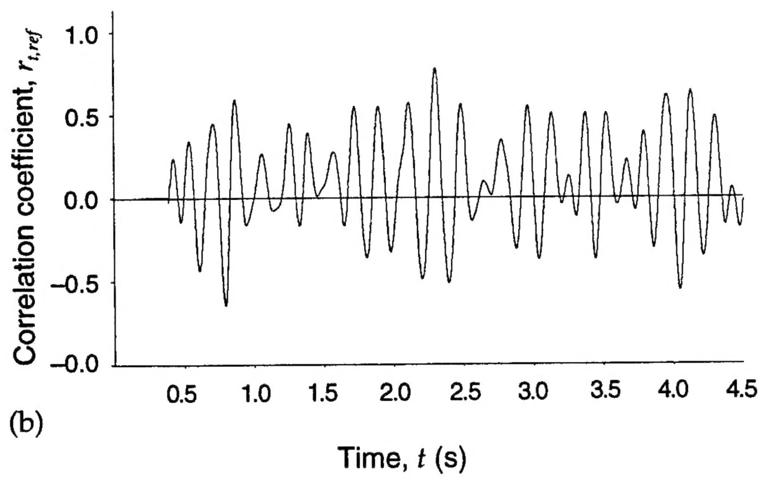
$$r_{t,\,ref} = \frac{\overline{u_t u_{ref}}}{\sqrt{\overline{u_t^2} \cdot \overline{u_{ref}^2}}}\,, \tag{1}$$

where $r_{t,ref}$ is the correlation coefficient, u_t is the test sine function $\sin(2\pi a t)$, and u_{ref} is the reference velocity. The test function is correlated with a segment of data of equal length, T, and the correlation coefficient is calculated. The test signal is then correlated with the data that have been shifted by some Δt, which is very small. If the data are made up of a series of data points, such as digital data, then the shift consists of one point. Thus, the entire original set (except the very first point), plus a new point at the end of the set, is correlated with the test function, and a new correlation coefficient is calculated. Eventually, as the reference velocity signal is advanced and correlated, a time history of $r_{t,ref}$ is created, as in figure 3b. This history of $r_{t,ref}$ enables us to accomplish two objectives: the detection of the

Figure 3. Reference roller velocities: (a) reference velocity fluctuations compared to sine function segment and (b) correlation coefficient with signal segment.



(a)



(b)

desired segment of reference velocity, and the alignment of the data sample, as the correlation coefficient will reach a peak and then reduce as the roller passes on. The data would then be collected for inclusion into the conditional averaging.

The likelihood of encountering a correlation of $r = 1.0$ is very small. In the present study, and as evident from the data sample in figure 3b, this never occurred. A threshold for r had to be established to provide enough conditional samples of the reference velocity to create a good statistical population. Since there was a band of frequencies in the reference velocities, the optimal frequency for the sine function also had to be picked.

3. Determination of Correct Test Frequency

The object in selecting a test frequency is to capture as many realistic ensembles as possible from a data set. The test frequency may initially be determined from the power spectrum plot of the reference velocity from figure 4. The frequencies ranged from 2 to 10 Hz, with a peak at about 6 Hz. This indicates that the energy distribution has a concentration in a periodic of that frequency. To determine the optimal frequency for the test signal, I performed a series of tests. The reference velocity data were correlated with a variety of sine functions with different values of a. The threshold correlation coefficient r_{th} was held to 0.5. This threshold coefficient seemed appropriate. The sample spanned 4000 points, or 0.4 s of data. This sample size was used since it was desired to capture at least one complete roller, plus most of the data associated with the rollers preceding and following the central roller. As each test frequency was used, the total number of ensembles captured, N, was recorded. Figure 5 shows the plot of N to a for the first experiment reference velocity. The peak number of captures appeared to be at about a = 8.5, which was different from the peak frequency in figure 4.

Figure 4. Linear plot of reference velocity power spectrum.

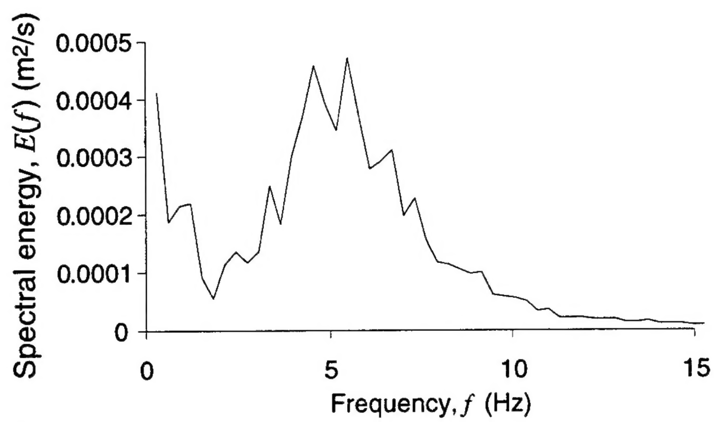
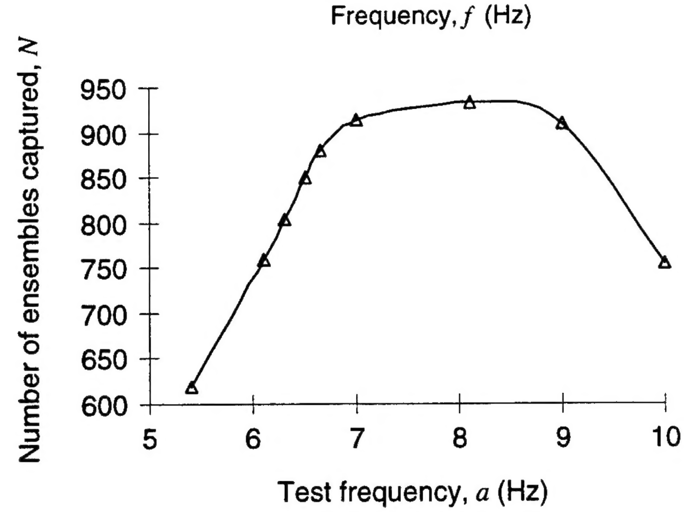


Figure 5. Number of ensembles captured; r_{th} = 0.5.

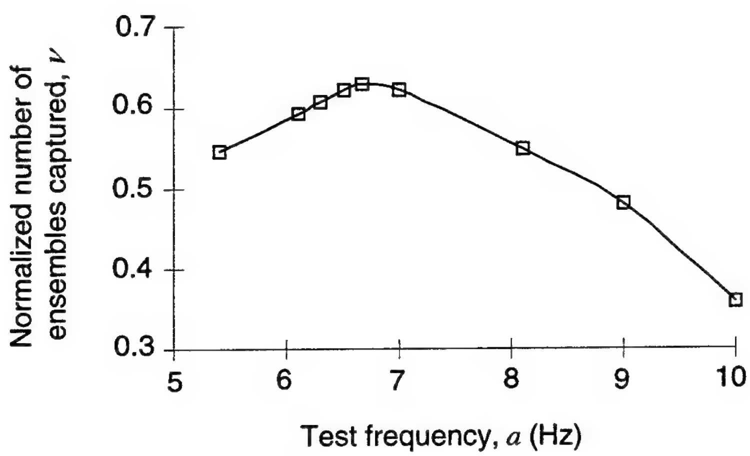


The large number of ensembles accumulated at a frequency other than at the power spectrum peak is due to the value used for the correlation threshold r_{th}. The value of the threshold would have a direct impact on the reference velocity bandwidth. A rudimentary method of eliminating the impact of r_{th} would be to normalize N by the total number of expected roller passages in the data set, if u_{ref} behaved as u_t. The expected number of ensembles for u_{ref} would be aT_{set}, where T_{set} is the time length of the entire data set. Normalization would be as

$$\nu = \frac{N}{aT_{set}} \quad . \tag{2}$$

Figure 6 shows the effect of this normalization. The peak is shifted back to the value indicated by the power spectrum. This indicates that looking at N alone will not enable selection of the appropriate test frequency. Care must be given to selecting the proper threshold correlation coefficient r_{th}.

Figure 6. Effect of normalization on test frequency.



4. Selection of Threshold Correlation Coefficient

Selection of the threshold correlation coefficient r_{th} affects the number of roller passages that will be included in the conditional average. Raising or lowering this value is, in effect, the same as changing the bandwidth of a band-pass filter. The bandwidth would be centered about the test frequency a, and rollers with associated reference velocity that have passed the frequency bandwidth would be included for conditional averaging. To estimate the bandwidth as a function of r_{th}, I compared the test signal with another function, $\sin(2\pi bt)$. The variance of a sine function is known: $(0.5)^{-1/2}$. Therefore, the correlation function for the continuous functions is modified from equation (1) to

$$r(a,b,T) = \frac{\frac{1}{T}\int_0^T \sin(2\pi at)\sin(2\pi bt)\,dt}{0.5} \quad . \tag{3}$$

This evaluates to

$$r(a,b,T) = \frac{2}{T}\left\{\frac{1}{2}\frac{\sin[2\pi T(a-b)]}{a-b} + \frac{1}{2}\frac{\sin[2\pi T(a+b)]}{a+b}\right\} \quad . \tag{4}$$

If we normalize b and the sample length T by the test signal frequency, which is a constant a, we obtain the response function in normalized terms:

$$\begin{aligned} \Theta &= aT \ , \\ \beta &= \frac{b}{a} \ , \\ r(\beta,\Theta) &= \frac{1}{\Theta}\left\{\frac{\sin[2\pi\Theta(1-\beta)]}{1-\beta} + \frac{\sin[2\pi\Theta(1+\beta)]}{1+\beta}\right\} \quad . \end{aligned} \tag{5}$$

Equation (5) can be used to determine the bandwidth of accepted normalized frequencies about $\beta = 1$ for a given r_{th}. It also shows that bandwidth is not solely dependent on frequency, but also on the size of the sample space T. Figure 7 shows the contour plot of r using the normalized variables. The reference frequency β is varied along the vertical, while Θ is varied along the horizontal. The contours are values of r.

Use of equation (5) helps to select the sample length and the associated threshold correlation coefficient. Figure 8 shows an example of bandwidth from given values of Θ and r.

With the threshold correlation coefficient set to 0.6, the range of frequencies passed is $\beta = 0.899$ to 1.096. Figure 9 shows the $\Delta\beta$ bandpass range with respect to the threshold correlation coefficient.

Figure 7. Distribution of r over β and Θ: (a) contour plots of r, (b) section plots of r at $\Theta = 1.5$.

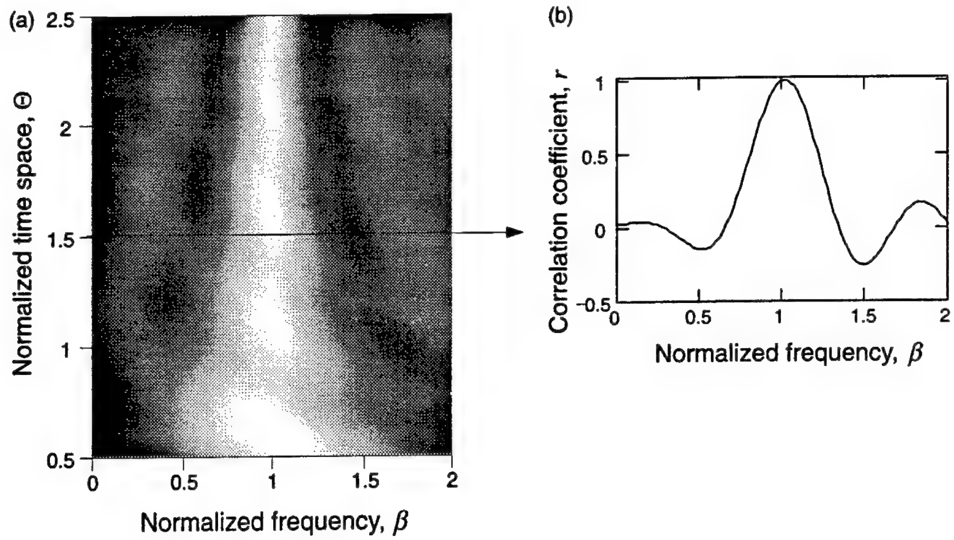


Figure 8. Bandwidth from given parameters.

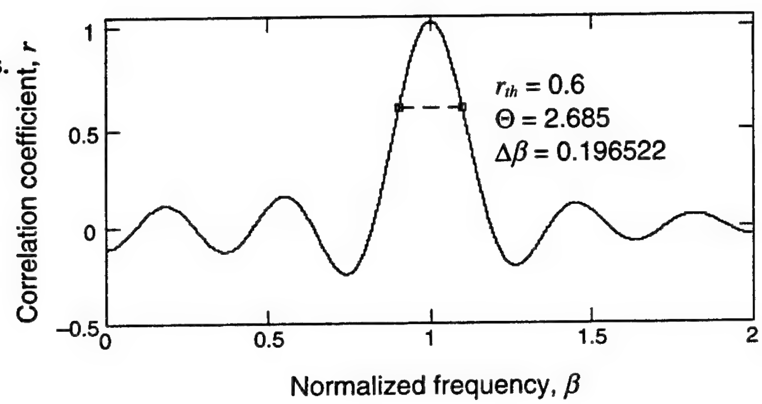
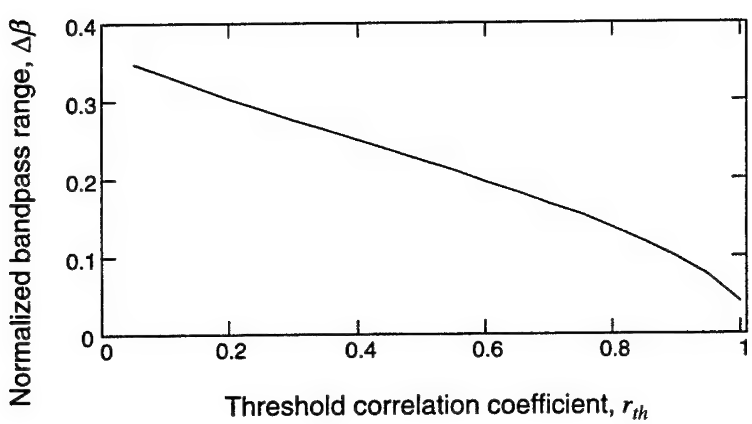


Figure 9. Bandpass range for $\Theta = 2.685$.



We still need to determine the appropriate test frequency, threshold correlation coefficient, and sample space. In the example of figures 8 and 9, a sample space of 4000 points was used (0.4 s of data), with the test frequency at 6.7 Hz, which is the peak value derived from figure 6. The bandwidth at $r_{th} = 0.6$ is $\Delta f = 1.32$ Hz, or a frequency range of 6.02 to 7.34 Hz. In figure 4, the center of the region where the roller passages contain the greatest spectral energy is at about 6 Hz. This high-energy region ranges from approximately 4 to 7 Hz. Assuming the time scale is smaller, fixed at $T = 0.2$ s, we can determine the approximate Δf from the spectral plot, and evaluate the lowest value of r_{th}. Assuming that the desired range would be 3 Hz, centered on 6 Hz, $\Delta\beta$ is then 0.5, or $\beta = 1.0 \pm 0.25$, and $\Theta = 1.2$. From figure 7(a) or equation (5), it is found that r_{th} should not be lower than 0.497. This threshold value would result in a bandpass of $\Delta\beta = 0.25$, so that the correlation-based conditional averaging would conservatively capture the desired information, while filtering out information with spectral energy outside the desired bounds.

5. Conclusion

The method of roller passage detection used by Hussain and Zaman[3] is frequent in the literature. Few of the methods described include any process to discriminate the types of rollers that pass (Subramanian et al[1]). In the self-similar region of the mixing layer, rollers sometimes merge or pair. The evolution of roller pairing involves a period in which the dissipation rate and turbulent kinetic energy levels are higher than during the period in which a roller is simply a pure roller. Inclusion of these types of activities into the conditional average would skew the distributions away from that of a pure roller. Use of the correlation coefficient effectively filters out roller pairings and rollers that may not conform to what may be considered normal. Conversely, the correlation coefficient can be used to exclude pure rollers, so that one can concentrate on those that are about to pair. More work needs to be done with this concept.

[1]*C. S. Subramanian, S. Rajagopalan, R. A. Antonia, and A. J. Chambers, "Comparison of conditional sampling and averaging techniques in a turbulent boundary layer," J. Fluid Mech.* ***123*** *(1982), 335–362.*

[3]*A.K.F.M. Hussain and K.B.M.Q. Zaman, "An experimental study of organized motion in the turbulent plane mixing layer," J. Fluid Mech.* ***159*** *(1985), 85–104.*

Distribution

Admnstr
Defns Techl Info Ctr
Attn DTIC-OCP
8725 John J Kingman Rd Ste 0944
FT Belvoir VA 22060-6218

Central Intellgnc Agency
Dir DB Standard
Attn GE 47 QB
Washington DC 20505

Chairman Joint Chiefs of Staff
Attn J5 R&D Div
Washington DC 20301

Defns Intllgnc Acgy
Attn DT 2 Wpns & Sys Div
Washington DC 20301

Dir of Defns Rsrch & Engrg
Attn DD TWP
Attn Engrg
Washintgon DC 20301

OIR CSB CRB
Attn A M Jones
RB 1413 OHM
Washington DC 20505

US Dept of Energy
Attn KK 22 K Sisson
Attn Tech Lib
Washington DC 20585

Commanding Officer
Attn NMCB23
6205 Stuart Rd Ste 101
FT Belvoir VA 22060-5275

DARPA
Attn Techl Lib
3701 N Fairfax Dr
Arlington VA 22203-1714

Dir of Chem & Nuc Ops DA DCSOPS
Attn Techl Lib
Washington DC 20310

Dpty Assist Scy for Rsrch & Techl
Attn SARD-TT F Milton Rm 3E479
Attn SARD-TT C Nash Rm 3E479

Dpty Assist Scy for Rsrch & Techl (cont'd)
Attn SARD-TT K Kominos
Attn SARD-TR R Chait
The Pentagon Rm 3E476
Washington DC 20310-0103

Hdqtrs Dept of the Army
Attn DAMO-FDQ D Schmidt
400 Army Pentagon
Washington DC 20310-0460

US Army Engrg Div
Attn HNDED FD
PO Box 1500
Huntsville AL 35807

US Army Mis & Spc Intllgnc Ctr
Attn AIAMS YDL
Redstone Arsenal AL 35898-5500

US Army NGIC
Attn Rsrch & Data Branch
220 7th Stret NE
Charlottesville VA 22901-5396

US Army Nuc & Cheml Agency
7150 Heller Loop Ste 101
Springfield VA 22150-3198

US Army Rsrch Ofc
Attn SLCRO-D
PO Box 12211
Research Triangle Park NC 27709-2211

US Army Strtgc Defns Cmnd
Attn CSSD H MPL Techl Lib
Attn CSSD H XM Dr Davies
PO Box 1500
Huntsville AL 35807

US Military Academy
Dept of Mathematical Sci
Attn MAJ D Engen
West Point NY 10996

Dept of the Navy
Chief of Nav OPS
Attn OP 03EG
Washington DC 20350

Nav Civil Engrg Lab
Attn Code L51 J Tancreto
Port Hueneme CA 93043-5003

Officer in Charge
Nav Construction Battlion Ctr
Civil Engrg Lab
Attn Code L31 Techl Lib
Port Hueneme CA 93041

Nav Rsrch Lab
Attn Code 2027 Techl Lib
Washington DC 20375

Ofc of Nav Rsrch
Attn Code 23 A Faulstick
800 N Quincy Stret
Arlington VA 22217

US Nav Academy
Attn Tech Lib
Attn LT D Fuller
Attn P McCoy
572 Holloway Rd
Annapolis MD 21402-5002

US Nav Academy
Attn R Smith
363 Rickover Hall 590 Holloway Rd
Annapolis MD 21402-5002

MIT
Attn Tech Lib
Cambridge MA 02139

The Johns Hopkins Univ
Applied Physic Lab
Attn Tech Lib
Johns Hopkins Rd
Laurel MD 20707

Univ of Maryland
Attn E Magrab
Attn Engrg & Techl Lib
Attn J Duncan
Attn J Wallace
Attn K Keiger
Attn L Ong

Univ of Maryland (cont'd)
Attn P Bernard
Attn R Calabrese
Rm 2168 Engrg Classroom Bldg
College Park MD 20742-5121

Olin Ordnance Product Material Control
Attn Tech Lib
Attn J Kibiger
10101 9th Stret N
ST Petersburg FL 33716

US Army Rsrch Lab
Attn AMSRL-IS-CS A Mark
Attn AMSRL-WT-TB J Condon
Attn AMSRL-WT-TB J Sullivan
Attn AMSRL-WT-TB P Muller
Attn AMSRL-WT-TB R Lottero
Attn AMSRL-WT-TC De Rosset
Attn AMSRL-WM-TA A Milhalcin
Aberdeen Proving Ground MD 21005-5055

US Army Rsrch Lab
Attn AMSRL-CI-LL Tech Lib (3 copies)
Attn AMSRL-CS-AL-TA Mail & Records Mgmt
Attn AMSRL-CS-AL-TP Techl Pub (3 copies)
Attn AMSRL-IS-EE A Wetmore
Attn AMSRL-IS-EE D Garvey
Attn AMSRL-IS-EE J Martin
Attn AMSRL-IS-EE K Deacon
Attn AMSRL-IS-EE R Cianco
Attn AMSRL-IS-EE R Loucks (4 copies)
Attn AMSRL-IS-EE R Meyers
Attn AMSRL-IS-EE R Pinnick
Attn AMSRL-IS-EE S Niles
Attn AMSRL-IS-ES P Gillespie
Attn AMSRL-SE-EA J Price
Attn AMSRL-SE-SA J Gerber
Attn AMSRL-SE-SA K Tran
Attn AMSRL-SE-SA M Fong
Attn AMSRL-SE-SA T Pham
Adelphi MD 20783-1197

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
April 1997

3. REPORT TYPE AND DATES COVERED
Final, from February to March 1997

4. TITLE AND SUBTITLE
Use of a Correlation Coefficient for Conditional Averaging

5. FUNDING NUMBERS
PE: 61102A

6. AUTHOR(S)
Richard B. Loucks

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
2800 Powder Mill Road
Attn: AMSRL-IS-EE
Adelphi, MD 20783-1197

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TN-91

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
2800 Powder Mill Road
Adelphi, MD 20783-1197

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
AMS code: 611102.53A
ARL PR: 7FEJ60

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A method of collecting ensembles for conditional averaging is presented that uses data collected from a plane mixing layer. The correlation coefficient of a sine function and a reference signal were used to determine the adequacy of blocks of the reference signal and the phase alignment for the data. Selection of the sine function period and a correlation coefficient threshold are discussed. Also examined are the effects of the period and threshold level on the number of ensembles captured for inclusion for conditional averaging. Both the selection of threshold correlation coefficient and the choice of data set length determine bandpass filtering; this study examined the effect of these choices on the ability to discriminate reference velocity ensembles for conditional averaging.

14. SUBJECT TERMS
Fluid dynamics, averaging, coherent flow, turbulence, conditional averaging

15. NUMBER OF PAGES
18

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102